

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-075968

(43)Date of publication of application : 23.03.2001

(51)Int.Cl. G06F 17/30
 G06F 12/00
 G10L 13/00
 G10L 15/08
 G10L 15/10
 G10L 15/18
 G10L 15/24

(21)Application number : 11-245535

(71)Applicant : TOSHIBA CORP

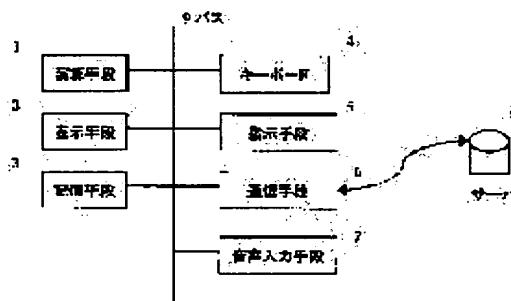
(22)Date of filing : 31.08.1999

(72)Inventor : SUGIKAWA AKIHIKO

(54) INFORMATION RETRIEVING METHOD AND RECORDING MEDIUM RECORDING THE SAME**(57)Abstract:**

PROBLEM TO BE SOLVED: To provide a recording medium recording an information retrieving method based on voice control with which the labor for an instruction to be applied to an information processor by a user is remarkably reduced, a using method is simplified and the time to be spent for acquiring desired information is remarkably reduced.

SOLUTION: This recording medium is provided as a program to operate on the information processor composed of an arithmetic means 1, a display means 2, a storage means 3, a communication means 6, a keyboard 4, an instruction means 5 and a voice input means 7. Then, the access instruction of a server is prepared while using code data for access provided from the plural voice recognized results and a template, By directly accessing a target Web page, the number of times of access to the server is reduced into one as well. As a result, even without knowing route information to the target Web page, the desired information can be acquired.

**LEGAL STATUS**

[Date of request for examination] 22.02.2002

[Date of sending the examiner's decision of rejection] 19.07.2005

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開2001-75968

(P2001-75968A)

(43)公開日 平成13年3月23日(2001.3.23)

(51)Int.Cl. ⁷	識別記号	F I	テーマコード(参考)	
G 0 6 F 17/30		G 0 6 F 15/403	3 1 0 Z	5 B 0 7 5
	12/00		5 4 6 T	5 B 0 8 2
G 1 0 L 13/00		15/40	3 1 0 F	5 D 0 1 5
	15/08	15/401	3 1 0 D	5 D 0 4 5
	15/10	G 1 0 L 3/00	R	
審査請求 未請求 請求項の数 8 O L (全 14 頁) 最終頁に続く				

(21)出願番号 特願平11-245535

(22)出願日 平成11年8月31日(1999.8.31)

(71)出願人 000003078

株式会社東芝

神奈川県川崎市幸区堀川町72番地

(72)発明者 杉川 明彦

兵庫県神戸市東灘区本山南町8-6-26

株式会社東芝関西研究センター内

(74)代理人 100059225

弁理士 葛田 瑋子 (外1名)

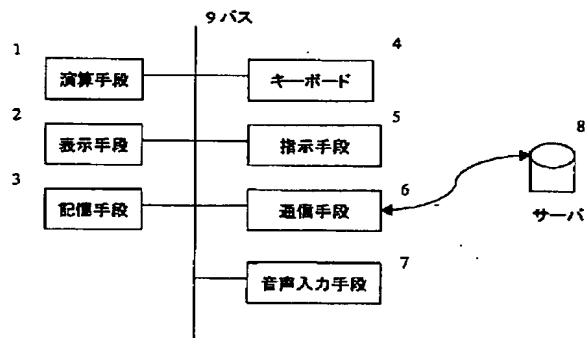
最終頁に続く

(54)【発明の名称】 情報検索方法及びそれを記録した記録媒体

(57)【要約】

【課題】 利用者が情報処理装置に与える指示の手間を大きく削減し、使用方法が容易で、かつ、所望の情報を取得するまでの時間を大幅に削減する、音声制御による情報検索方法を記録した記録媒体を提供する。

【解決手段】 演算手段1、表示手段2、記憶手段3、通信手段6、キーボード4、指示手段5、音声入力手段7から構成される情報処理装置上で動作するプログラムとして実現され、複数の音声認識結果から得られたアクセス用のコードデータと、テンプレートとを用いてサーバのアクセス命令を作成し、直接目的のWebページにアクセスすることで、サーバへのアクセス回数も1回だけとなる。この結果、目的のWebページに至る経路情報を知らなくても所望の情報を入手することができる。



1

【特許請求の範囲】

【請求項 1】電気通信回線を通じて、ハイパーテキスト方式によってサーバから提供される複数の表示単位から構成される集合情報から、利用者が必要とする情報を検索する方法であって、
前記利用者が前記必要とする情報に関して発声した音声の内容を認識する音声認識ステップと、
前記音声認識ステップにおいて認識した認識結果を所定の条件に基づいて分類する認識結果分類ステップと、
前記認識結果分類ステップにより分類した 1 つ以上の認識結果を利用して、前記集合情報から、前記必要とする情報を検索するためのアクセス命令を作成するアクセス命令作成ステップと、
前記アクセス命令作成ステップにおいて作成したアクセス命令を、前記電気通信回線を通じて前記サーバへ送信する送信ステップと、
前記検索した情報を、前記電気通信回線を通じて前記サーバから受信する受信ステップと、
前記受信ステップにおいて受信した情報の一部または全体を出力する出力ステップと、
を有することを特徴とする情報検索方法。

【請求項 2】前記受信ステップにおいて受信した情報から、前記認識結果に関連する情報を選択する関連情報選択ステップを有することを特徴とする請求項 1 記載の情報検索方法。

【請求項 3】前記アクセス命令作成ステップは、前記認識結果分類ステップにより認識した 1 つ以上の認識結果を利用して複数のアクセス命令を作成することを特徴とする請求項 1 記載の情報検索方法。

【請求項 4】前記認識結果分類ステップにより分類した認識結果を、その分類に対応させた状態で前記利用者に表示する表示ステップを有することを特徴とする請求項 1、2、3 記載の情報検索方法。

【請求項 5】前記利用者の状況を認識する状況認識ステップを有し、
前記アクセス命令作成ステップは、
前記認識結果分類ステップにより分類した前記 1 つ以上の音声に基づく認識結果と、前記状況認識ステップにおいて認識した 1 つ以上の状況に基づく認識結果を利用して前記アクセス命令を作成することを特徴とする請求項 1 記載の情報検索方法。

【請求項 6】前記状況認識ステップで認識する利用者の状況以外の他の状態を認識する認識入力ステップと、
前記アクセス命令作成ステップは、
前記認識結果分類ステップにより分類した前記 1 つ以上の音声に基づく認識結果と、前記認識入力ステップにおいて認識した他の状態に基づく認識結果を利用して前記アクセス命令を作成することを特徴とする請求項 5 に記載の情報検索方法。

【請求項 7】電気通信回線を通じて、ハイパーテキスト

2

方式によってサーバから提供される複数の表示単位から構成される集合情報から、利用者が必要とする情報を検索する方法を実現するためのプログラムを記録した記録媒体であって、

前記利用者が前記必要とする情報に関して発声した音声の内容を認識する音声認識機能と、

前記音声認識機能において認識した認識結果を所定の条件に基づいて分類する認識結果分類機能と、

前記認識結果分類機能により分類した 1 つ以上の認識結果を利用して、前記集合情報から、前記必要とする情報を検索するためのアクセス命令を作成するアクセス命令作成機能と、

前記アクセス命令作成機能において作成したアクセス命令を、前記電気通信回線を通じて前記サーバへ送信する送信機能と、

前記検索した情報を、前記電気通信回線を通じて前記サーバから受信する受信機能と、

前記受信機能において受信した情報の一部または全体を出力する出力機能と、

を実現するプログラムを記録したことを特徴とする情報検索方法の記録媒体。

【請求項 8】電気通信回線を通じて、ハイパーテキスト方式によってサーバから提供される複数の表示単位から構成される集合情報から、利用者が必要とする情報を検索する装置であって、

前記利用者が前記必要とする情報に関して発声した音声の内容を認識する音声認識手段と、

前記音声認識手段において認識した認識結果を所定の条件に基づいて分類する認識結果分類手段と、

前記認識結果分類手段により分類した 1 つ以上の認識結果を利用して、前記集合情報から、前記必要とする情報を検索するためのアクセス命令を作成するアクセス命令作成手段と、

前記アクセス命令作成手段において作成したアクセス命令を、前記電気通信回線を通じて前記サーバへ送信する送信手段と、

前記検索した情報を、前記電気通信回線を通じて前記サーバから受信する受信手段と、

前記受信手段において受信した情報の一部または全体を出力する出力手段と、

を有することを特徴とする情報検索装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、例えば、Web サーバの提供する情報へのアクセスを実現するための情報検索方法及びそれを記録した記録媒体に関する。

【0002】

【従来の技術】近年、インターネット上に、時刻表、番組表、天気、ニュースなどさまざまな情報が提供されている。利用者は、携帯型の情報処理装置上で、情報アク

50

3

セス用のアプリケーションであるブラウザを使用して、任意の場所から必要な情報にアクセス可能となった。利用者はなんらかの情報を知りたい場合には、最初にポータルと呼ばれる起点となるWebページにアクセスし、前記Webページから関連するページを順次指示することで目的の情報を得ることができる。

【0003】一般的にブラウザは、受信したHTML (Hyper Text Markup Language) 文書を解析し、表示用データを作成し画面に提示する。このとき画面サイズが文書サイズより小さい場合は、文書の一部分のみ表示し、利用者は必要に応じて表示画面のスクロールを指示する。Webページによっては、1つのページに多くの情報を含んでいる場合があり、利用者は目的とする情報を知るためには、前記当する場所までページをスクロールする必要がある。特に携帯型の情報処理装置では、表示画面が小さいために移動量は大きくなる問題が生じる。

【0004】一方、近年、携帯型情報処理装置の普及により、多くの利用者がオフィスや自宅以外の様々な場所で利用する機会が増えている。これまでのように机の前でイスに座りながらキーボードやマウスで操作するスタイルだけではなく、場所・用途に適した入力方法が必要になってきている。

【0005】音声入力は、ハードウェアの進歩と共に実用化され始めており、メニュー選択や、文書入力手段として利用されている。最近では音声制御で、WWW (World Wide Web) 上のWebページにアクセスする方法も実現されている。例えば、予め登録されたキーワード発話により、そのキーワードに対応したWebページへのアクセスを実現した音声制御ブラウザが利用されている。この方法では、ポータルと呼ばれる他のWebページへのリンクを集めたWebページには、音声でアクセスできるが、そこからリンク先のページへの移動は、利用者がマウスやキーボードなどの他の手段で指示する必要がある。

【0006】このような問題に対処するために、アンカーと呼ばれるリンク先を示す文章そのものを発話する方法 (特開平10-124293号)、あるいは、アンカー内の文章から選択したキーワードを発話する方法 (特開平11-25098号) など、音声の指示だけで関連するWebページへの移動を実現する方法が提案されている。

【0007】上記提案では、受信後したHTMLで記述された文書を解析し、TAGと呼ばれる識別子を手がかりに、アンカー文とリンク先情報 (URL) を取得し、メモリに対応関係を一時的に記録し、音声辞書には、アンカー文の読みデータ、あるいはアンカー文から抽出したキーワードの読みデータを登録する。音声認識エンジンが利用者の発話を認識すると、システムは認識結果に対応するリンク先情報 (URL) を取得し、ブラウザに

4

そのURLのWebページにアクセスする命令を与える。ブラウザはWebページ取得後、システムは再びHTML文章を解析することで更なる移動を実現する。

【0008】このように音声だけで、目的とするWebページまでアクセス可能となったが、サーバが提供する情報が多い場合には、そのWebページの構成を把握していないと、目的のWebページへの到達までに試行錯誤を繰り返すこととなる。また、目的のWebページまでの到達方法を把握していても、何回もの指示を必要とする。それゆえ、必要とする情報の取得には手間や時間がかかり、利用者の利便性を損ねている。

【0009】特に、Webページを順番にたどっていく場合など、通信回線の状況に遅延時間は左右されやすく、またWebページのデータ量が大きい場合には、数分から十数分のアクセス時間が必要となる。

【0010】音声認識技術は、完全ではなく、時には特定キーワードの認識ができない場合がある。そのような場合のため何らかの代替手段は必要である。また、予め認識対象をエンジンに登録しておく必要があり、対象とする件数が多い場合には、認識に要するメモリサイズや遅延時間など、携帯情報処理装置を考慮すると実用上の課題が生じる。

【0011】条件を変更して同じサーバに再度検索する場合には、一旦検索条件を設定するページに戻り、再度入力を行う必要があり、これも利用者の利便性を損ねている。従来、複数の種類のデータを音声で入力する場合には、銀行のATMのように予め決められた手順に従って入力を行う必要があり、システム主導で対話が進んでいく。

【0012】それに対して、図14のようなデータ入力画面を利用者に提示し、利用者にデータ入力を行うスロットを選択させ、そのスロットに関連した認識辞書を準備して、データ入力を行う方法が考えられる。この方法ではスロットに関連したデータしか認識しないため、他のスロット向けの単語を誤認識することはない。また、誤認識が生じた場合でも、正しく認識するまで再度発話するだけで良い。しかしこの方法では、指示手段でスロットを指定する手間が生じる。

【0013】

【発明が解決しようとする課題】以上説明したように従来の音声制御ブラウザでは、必要な情報の所得までに要する手間や時間が多いという問題があった。

【0014】本発明は、利用者が情報処理装置に与える指示の手間を大きく削減し、使用方法が容易で、かつ、所望の情報を取得するまでの時間を大幅に削減する、音声制御による情報検索方法及びそれを記録した記録媒体を提供することを目的とする。

【0015】

【課題を解決するための手段】請求項1の発明は、電気通信回線を通じて、ハイパーテキスト方式によってサー

5

バから提供される複数の表示単位から構成される集合情報から、利用者が必要とする情報を検索する方法であって、前記利用者が前記必要とする情報に関して発声した音声の内容を認識する音声認識ステップと、前記音声認識ステップにおいて認識した認識結果を所定の条件に基づいて分類する認識結果分類ステップと、前記認識結果分類ステップにより分類した1つ以上の認識結果を利用して、前記集合情報から、前記必要とする情報を検索するためのアクセス命令を作成するアクセス命令作成ステップと、前記アクセス命令作成ステップにおいて作成したアクセス命令を、前記電気通信回線を通じて前記サーバへ送信する送信ステップと、前記検索した情報を、前記電気通信回線を通じて前記サーバから受信する受信ステップと、前記受信ステップにおいて受信した情報の一部または全体を出力する出力ステップと、を有することを特徴とする情報検索方法である。

【0016】請求項2の発明は、前記受信ステップにおいて受信した情報から、前記認識結果に関連する情報を選択する関連情報選択ステップを有することを特徴とする請求項1記載の情報検索方法である。

【0017】請求項3の発明は、前記アクセス命令作成ステップは、前記認識結果分類ステップにより認識した1つ以上の認識結果を利用して複数のアクセス命令を作成することを特徴とする請求項1記載の情報検索方法である。

【0018】請求項4の発明は、前記認識結果分類ステップにより分類した認識結果を、その分類に対応させた状態で前記利用者に表示する表示ステップを有することを特徴とする請求項1、2、3記載の情報検索方法である。

【0019】請求項5の発明は、前記状況認識ステップで認識する利用者の状況を認識する状況認識ステップを有し、前記アクセス命令作成ステップは、前記認識結果分類ステップにより分類した前記1つ以上の音声に基づく認識結果と、前記状況認識ステップにおいて認識した1つ以上の状況に基づく認識結果を利用して前記アクセス命令を作成することを特徴とする請求項1記載の情報検索方法である。

【0020】請求項6の発明は、前記利用者の状況以外の他の状態を認識する認識入力ステップと、前記アクセス命令作成ステップは、前記認識結果分類ステップにより分類した前記1つ以上の音声に基づく認識結果と、前記認識入力ステップにおいて認識した他の状態に基づく認識結果を利用して前記アクセス命令を作成することを特徴とする請求項1に記載の情報検索方法である。

【0021】請求項7の発明は、電気通信回線を通じて、ハイパーテキスト方式によってサーバから提供される複数の表示単位から構成される集合情報から、利用者が必要とする情報を検索する方法を実現するためのプログラムを記録した記録媒体であって、前記利用者が前記

6

必要とする情報に関して発声した音声の内容を認識する音声認識機能と、前記音声認識機能において認識した認識結果を所定の条件に基づいて分類する認識結果分類機能と、前記認識結果分類機能により分類した1つ以上の認識結果を利用して、前記集合情報から、前記必要とする情報を検索するためのアクセス命令を作成するアクセス命令作成機能と、前記アクセス命令作成機能において作成したアクセス命令を、前記電気通信回線を通じて前記サーバへ送信する送信機能と、前記検索した情報を、前記電気通信回線を通じて前記サーバから受信する受信機能と、前記受信機能において受信した情報の一部または全体を出力する出力機能と、を実現するプログラムを記録したことを特徴とする情報検索方法の記録媒体である。

【0022】請求項8の発明は、電気通信回線を通じて、ハイパーテキスト方式によってサーバから提供される複数の表示単位から構成される集合情報から、利用者が必要とする情報を検索する装置であって、前記利用者が前記必要とする情報に関して発声した音声の内容を認識する音声認識手段と、前記音声認識手段において認識した認識結果を所定の条件に基づいて分類する認識結果分類手段と、前記認識結果分類手段により分類した1つ以上の認識結果を利用して、前記集合情報から、前記必要とする情報を検索するためのアクセス命令を作成するアクセス命令作成手段と、前記アクセス命令作成手段において作成したアクセス命令を、前記電気通信回線を通じて前記サーバへ送信する送信手段と、前記検索した情報を、前記電気通信回線を通じて前記サーバから受信する受信手段と、前記受信手段において受信した情報の一部または全体を出力する出力手段と、を有することを特徴とする情報検索装置である。

【0023】請求項1、2、3、4、7、8に係る発明によれば、利用者は複数のキーワードの発話、あるいは文章を発話するだけで、ハイパーテキスト方式によってサーバから提供される複数の表示単位から構成される集合情報から、利用者が必要とする情報を検索する場合に、大幅に手間とアクセスまでの遅延時間を減少させることが可能となる。

【0024】また、利用者の望む情報のみ表示することで、スクロールなどの手間を省くことが可能となり、小型の情報処理装置での利便性が良くなる。

【0025】さらに、利用者主導での入力方式を提供することにより利便性を向上しつつ、入力時の手間を削減している。

【0026】請求項5、6に係る発明によれば、音声以外の認識手段と組み合わせることで、入力の手間の削減や音声入力ができない場合でも所望の情報にアクセス可能となる。

【0027】

【発明の実施の形態】以下、本発明の実施形態を図面に

基づいて説明する。

【0028】第1の実施例

本発明の第1の実施例を、図1から図8を用いて説明する。

【0029】図1は、情報処理装置の構成を示すブロック図の一例であり、演算手段1、表示手段2、記憶手段3、通信手段6、キーボード4、指示手段5、音声入力手段7から構成される情報処理装置上で動作するプログラムとして実現する。

【0030】また、CD-ROMやFDなどの記録媒体、あるいは通信手段6を通じてこの情報処理装置に読み込まれるものとする。

【0031】記憶手段3、一時的なデータの保持に使用されるメモリと呼ばれるRAM、永久に保存するためのハードディスクなどの装置を用いる。RAMには本プログラムや認識用辞書データやリンク管理テーブルなどの作業用の変数が格納され、ハードディスクは、受信したHTML文章のキャッシュ、音声認識用の音素データ、読みデータ辞書などを記録する。

【0032】表示手段2は、表示内容を一時的に記憶するビデオメモリ、DA変換部、CRTや液晶表示装置などで構成する。

【0033】通信手段6は、ケーブルを使用する有線LAN、RS232C、セントロニクス、SCSI、USB、IEEE1394など、あるいはケーブルを使用しない、無線LAN、赤外線通信、PHSデータ通信などで実現する。

【0034】指示手段5は、マウス、トラックボール、アキュポイント、ペン入力などで実現する。

【0035】音声入力手段7は、マイクや外部アンプと接続し、入力された音声信号を、AD変換を通じてデジタル信号に変換する。入力されたデータはメモリに格納する。

【0036】演算手段1は、CPUとも呼ばれ上記の各手段とバス9を通じて接続されデータの送受信を行う。

【0037】HTML文章を解析し画面に表示を行う処理は、コンポーネント化されたブラウザモジュールを使用することにより可能である。例えばマイクロソフト社のブラウザは容易に他のプログラムからその機能を利用できる。

【0038】(状況の説明) 本実施例が有効に機能する状況を説明するため、タスクを駅施設の検索と仮定する。

【0039】このような情報を提供するWebページとしては、「駅前探検クラブ(商標、以下同様)」、<http://ekimae.toshiba.co.jp/>などが存在する(平成11年8月17日現在)。利用者は、このWebページを通じて、利用者が東京駅のタクシー乗り場に関する情報を音声指示で取得することと設定する。

【0040】(従来のアクセス方式) 最初に従来方式でのアクセスを説明する。

【0041】利用者は、駅前探検クラブへのポインターをブラウザに記録しておき、ホームページやポータルと呼ばれる起点となるWebページにアクセスする。

【0042】次に利用者はタスクを指定して、2番目に路線名を指定、さらに駅名を指定、最後に施設を指定するよう順番に検索用キーを指定することで、最終的に所望のWebページにたどりつく。

10 【0043】前記Webページのサーバは、駅施設に関するデータベースを所持しており、CGI(Common Gateway Interface)と呼ばれる技術を用いて、利用者からの問い合わせに対し、データベースを検索し、問い合わせ毎に返答用のWebページを作成し、クライアントであるブラウザにHTML文章を送信する。

【0044】図2は、東京駅のタクシー乗り場を検索するため識別子の例であり、図3に示す11から14の4つの要素で構成される。

20 【0045】11は、URLと呼ばれる特定のWebページを示すものである。

【0046】12から14は、CGI用のパラメータである。

【0047】12は、路線を示すコードである。

【0048】13は、駅を示すコードである。

【0049】14は、施設を示すコードである。

【0050】サーバ上のデータベースアクセスプログラムは、これらを検索用のキーとしてデータベースを検索する。

30 【0051】上記のサーバへのアクセスの例では、利用者がWebページを順番にアクセスする毎にCGI用のコードが付与され、サーバは全てのコードが取得できた場合に、データベースの検索を行う。このときサーバは受信したコードの妥当性を検査し、正しければ各々のコードをキーにデータベースを検索し、その結果をクライアントに送信する。

40 【0052】この従来例では、URLにデータベース検索用のパラメータ付加してを送信しているが、他のパラメータ送信方法(例えば、異なるプロトコルを使用する)でも同等の効果を有する。

【0053】(本実施例のアクセス方式) 図4は、本実施例に係るフローチャートの一例である。

【0054】本プログラムは、図1の情報処理装置上で動作する。

【0055】1. ステップ101プログラムは、音声入力手段7からの定期的割り込みにより、一定量の音声データを音声入力手段7から読み出し、メモリに記憶する。

50 【0056】2. ステップ102そのメモリに記憶されたデータを用いて入力された音声データの類似度を計算

する。

【0057】類似度計算は、予め決められた条件を満たした入力データを音声データとみなし、そのデータにフーリエ変換を行い、周波数信号に変換し、特徴データを作成する。

【0058】また、辞書に登録した単語と音素モデルから各単語の音素モデルを作成し、入力データの特徴データと複数からなる認識単語データの特徴データとの距離計算を行い、各単語についてのスコアを求める。

【0059】計算結果が予め決められた閾値を超える場合には、最良のスコアが得られた単語を認識結果として出力する。

【0060】閾値に達しない場合には、エラーを出力する。認識結果がエラーの場合には、ステップ101に戻って音声入力待ちの状態となる。

【0061】認識辞書に登録する内容は、全ての路線名、全ての駅名、全ての施設名の読みデータである。

【0062】また、各読みデータに対して属性情報を付与したテーブルを作成し、予め記憶手段3に格納しておく。

【0063】図5はテーブルへの登録方法である。

【0064】読みデータに対して、表示用データ、種別、コードの並びになっている。同時にシステム制御用の単語も登録する。例では、検索開始のトリガーとなる単語を示す。同様の属性データと関連付けたテーブルを作成する。テーブルは、左から順に、読みデータ、表示データ、種別である。

【0065】3. ステップ103

ステップ102で認識が成功したならば、認識結果の単語、例えば「とうきょう」をキーとし、上述したテーブルを検索する。「とうきょう」のキーからは、表示用のデータ「東京」、入力されたデータの種別「駅」、データベース検索用のコード「0000100」が得られる。

【0066】4. ステップ104

得られたデータは、一旦メモリに記録する。

【0067】タスクによっては、検索用のコードは固定なものだけでなく、日時など動的に変化するものがある。

【0068】例えば、利用者が「明日」と発話したら、認識エンジンからは「あす」という情報が得られる。このような場合は、当日（仮に99年7月14日と仮定する）の情報を利用して、検索用コード990715を作成するステップを準備する。この場合、上記のテーブルは、データの種別に演算を、検索用のコードには日付計算用の演算識別子を入れておく。演算手段1には、検索用のコードと表示用のデータを作成するサブルーチンを準備しておく。

【0069】5. ステップ105

次に、プログラムは入力が終了したかどうか、判定す

る。判定方法は、データベースへのアクセス用に必要なコードが全てそろったか検査する。

【0070】上記の例では、路線コード、駅コード、施設コードの3つが入力されたか確認する。

【0071】この場合、誤認識の可能性もありうるので一定期間入力待ち、その期間内に更なる入力が行われない場合は、これまで入力されたデータを確定したものと判断する。

【0072】あるいは認識エンジン辞書にプログラム制御用のデータを登録しておき、入力された単語が制御命令を示す単語と一致したら検索開始する方法でも同様の効果が得られる。

【0073】あるいは、ボタンなどの指示手段5からの入力をトリガーに用いる方法もある。

【0074】6. ステップ106

アクセス命令作成ステップは、準備してあるテンプレートに、確定したコードを挿入しWeb上のサーバへのアクセス用データを作成する。

【0075】例えば、駅施設情報アクセス用には図6のようなURLをテンプレートとして準備しておく。

【0076】前ステップで確定したコードをその種別に従い、前記テンプレートの前記当する部分、例えば駅を示すコードであれば「ekiCode=」のあとに確定したコード「0000100」を挿入する。テンプレートはタスクに応じて予め記憶手段3に記録しておき、タスクが決定した時点で記憶手段3から取得する。

【0077】タスクの設定は、予め利用者が明示的に指示する。

【0078】あるいは、入力されたキーワードの種類データを利用し、自動的に決定する方法も考えられる。

【0079】例えば、路線名、駅名、施設名が入力されたら駅施設検索であることが決まるように、各々関連付けたテーブルを記憶手段3に予め準備することで実現できる。

【0080】7. ステップ107

アクセス送信命令ステップは、前ステップで作成したURLを用いて、通信手段6を通じて情報取得要求を送信する。

【0081】8. ステップ108

HTML文書受信ステップは、サーバから送られてきたHTML文書を一旦情報処理装置のメモリやキャッシュに保存する。

【0082】9. ステップ109

画面表示ステップは、受信した文書をTAG情報に基づき解析し、表示用データを作成し、表示手段2を通じて利用者に表示する。

【0083】ステップ107からステップ109は、ブラウザと呼ばれるモジュールを利用することにより容易に実現できる。

【0084】さらに、画面表示（ステップ109）の後

で、利用者は他の施設に関して再度アクセスする場合がある。

【0085】ここでは仮に「バス乗り場」と仮定する。利用者の施設名の発話により、ステップ101からステップ104を通じて、表示用データ、データの種別、検索用コードがメモリに記憶される。入力終了判定ステップ（ステップ106）では、前回入力された路線データ、駅データと新たに入力された施設データから入力終了と判定する。

【0086】以上要約すると、複数の音声認識結果から得られたアクセス用のコードデータと、テンプレートとを用いてサーバのアクセス命令を作成し、直接目的のWebページにアクセスすることで、サーバへのアクセス回数も1回だけとなる。この結果、目的のWebページに至る経路情報を知らなくても所望の情報を入手することができる。また従来、複数回のアクセスが必要であったが、その回数を1度に低減でき情報所得時の通信遅延時間を短くすることができる。

【0087】以上説明したように、本実施例により、利用者は必要な情報を取得する場合には、必要なキーワードを発話するだけでという容易な方法で、かつ入手までの遅延時間の少ない音声制御によるWebサーバへのインタフェースが実現できる。

【0088】（他のアクセス形式）上記の説明では、キーワードを複数回発話する方式を説明したが、音声認識技術の他の方式であるワードスポット方式でも実現できる。

【0089】すなわち、利用者の自由発話から検索に関連する1つ以上のキーワードを抽出し、それらを用いて前述のステップを実行することとなる。

【0090】このワードスポット方式では、ラティス解析に計算機資源が必要となるが、反面利用者は会話に近い形で質問を入力できる利点がある。

【0091】この方法を図7に示すフローチャートを用いて説明する。上記アクセス方式と重複する部分を除き、異なるステップのみ説明する。

【0092】1. ステップ111

利用者からの音声が入力される。

【0093】図8は、その利用者から入力された文章を示すものである。

【0094】2. ステップ112

認識エンジンは、単語辞書に登録された各データの類似度を計算し、閾値以上の単語に関して、認識結果とスコアや発話時刻情報を属性として出力する。通常このような単語の集まりをキーワードラティスと呼ぶ。

【0095】3. ステップ113

ワードスポット方式は、けん切りと呼ばれる単語と単語の切れ目をうまく検出できない場合があり、ラティスの中には不要な単語が多く含まれている。

【0096】それゆえ、ラティス解析ステップ（ステッ

プ113）では、予め登録された文法的制約、各候補のスコア、発話時刻情報を元に利用者の発話を推定する。

【0097】4. ステップ114

推定結果が決められた条件を満たしている場合は、推定した文の中に含まれるキーワードを取得する（ステップ114）。

【0098】基準に達しない場合は、その旨を利用者に通知し、再度利用者からの入力を待つ（ステップ111）。

【0099】以下のステップ115から118は、前述したステップ106から109と同じ処理を行うことにより、自然言語を用いてインターネット上のサーバから必要な情報が取得できる。

【0100】（変更例1）他の実施例として音声合成を出力手段に用いることも可能である。

【0101】この場合、表示手段3は必ずしも必要ではなく、与えられた文字データを音声データに変換する音声合成エンジンモジュールが必要となる。音声合成エンジンは、プログラムや、DSPを使用して実現可能である。音声認識及び音声合成を用いるインタフェースは、ハンズフリーで使用でき、両手がふさがっている状況でも必要な情報を得ることが可能となる。

【0102】（変更例2）本実施例には、駅施設情報のアクセスを例にしたが、駅施設に限定されるものではなく、旅行情報、株価、本、パソコンのパーツ、料理レシピ、医療情報、テレビ番組、ニュース、地図、中古車、不動産、オークションなど他の多くの分野にも容易に適応可能である。

【0103】（変更例3）検索エンジンと呼ばれるWebページにアクセスすることも考えられる。

【0104】一例としては、入力されたキーワードとANDの論理演算を組み合わせる検索式を作り、検索サービスを提供するWebページにアクセスし、検索結果を取得して提示する。

【0105】あるいは、ラティス解析を行うことにより、AND、OR、NOTなどの論理式を組み合わせる検索することも可能である。

【0106】（変更例4）本実施例では、音声認識手段を自装置に有する方式を説明したが、音声認識手段をサーバクライアント方式で構成する方法も可能である。

【0107】この場合、自装置の通信手段6を通じてサーバである他の情報処理装置とコネクションを確立し、入力された音声信号をデジタル化したデータそのもの、あるいは特徴データに変換した後のデータをサーバに送信し、認識結果を受信することにより実現する。

【0108】以上のように、変更したい条件のみ発話することで、前記情報を提供するWebページに一度でアクセスが可能となり、従来の方法と比較して大幅に手順や時間の短縮が実現できる。

13

【0109】第2の実施例

本発明の第2の実施例を図9から図11を用いて説明する。

【0110】第1の実施例と異なる点のみ説明するものとする。

【0111】また、本実施例の効果を示せるように第1の実施例と同様の状況を想定するが、本説明では、サーバから送信された情報(ステップ128)に複数の情報が含まれている場合を仮定する。

【0112】例えば、「タクシー乗り場」のほかに「バス乗り場」が含まれているものとする。

【0113】他の施設に関する情報も含んでいても、本実施例の説明には影響しないので、ここでは施設の数2つとする。

【0114】ステップ121から128は、第1の実施例のステップ101から108と同様である。

【0115】1. ステップ129

関連情報選択ステップでは、受信した文章を、TAGを元に解析し、利用者に必要な情報のみ選択する。

【0116】図10は選択した文字情報の例を示す。

【0117】入力データの分類ステップ(ステップ123)で得られた、施設を示す情報である「タクシー乗り場」をキーにして、図10に示す文字情報から、関連する情報のみを選択する。

【0118】2. ステップ130

画面表示ステップでは、選択したデータを元に図11に示す表示用文字データを作成し、表示手段3に出力する。

【0119】本例では「京浜東北」線を利用者は指示しているが、そのまま解釈すると前記当件数が0件になってしまう。そのような応答は、利用者の立場で考えると不便なものである。ゆえに前記当するものがない場合、次候補を提示する様にプログラムを作成しておく。もちろん選択したデータに、京浜東北線に係る情報がある場合は、施設同様に(ステップ123)で取得した表示データ「京浜東北」を手がかりに、前記当するデータのみ選択すれば良い。

【0120】以上説明した方法により、利用者の望む情報のみ表示することで、スクロールなどの手間を省くことが可能となり、小型の情報処理装置、例えば、時計、携帯電話、カードなどでの利用が可能となる。

【0121】第3の実施例

本発明の第3の実施例を図12のフローチャートを用いて説明する。

【0122】第1の実施例と異なる点のみ説明するものとする。

【0123】また、本実施例の効果を示せるように第1の実施例と同様の状況を想定するが、本実施例では、必要な情報に複数回のアクセスが必要なものと仮定する。

【0124】つまり、Webページの中に選択肢がアン

14

カーとして提示されており、アンカーを順番に選択していくことにより、必要なWebページにアクセスできるものと仮定する。ここではサービスの例として、路線名、駅名、施設名を順番に選択して、次々にWebページを更新するサーバを対象とする。

【0125】1. ステップ131から135

ステップ131からステップ135を通じて、検索に必要なキーワードはメモリに記録されているものとする。第3の実施例では図5に示すコードデータは使用せず、認識結果をもとに、表示データと種別データを対にして一時的にメモリに記録しておく。

【0126】2. ステップ136

また、指示手段5や音声入力手段7を通じてタスクは、すでに指定されているものとする。タスクに対応したWebページのURLは、予めタスクと対応させて記憶手段3に記憶させておき、それを検索することにより取得する。

【0127】3. ステップ137

取得したURLを用いてWebページアクセス命令を送信する。

【0128】4. ステップ138

次に通信手段6を通じてHTML文章を受信する。

【0129】5. ステップ139

HTML文章が、目的の文章かどうか判定する。

【0130】終了判定は、予め検索に必要なキーワードをスタックに格納しておき、キーワードを用いてアクセス命令を作成する毎に、スタックから前記キーワードを削除していき、残りのキーワードを計数することで実現できる。

【0131】6. ステップ140

受信終了でない場合は、HTML文章解析ステップに進み、受信したHTML文章を、TAGを手がかりにアンカーとリンク先識別子(URL)に分類し、その対応表を作成する。最初の路線名を選択するWebページの場合は、路線名を示すアンカーと路線に対応した駅を選択するWebページのURLが対になったデータが作成される。

【0132】7. ステップ136から140の繰り返し
次にスタックを検査し、種別が路線である情報、この例では「京浜東北」を取得し、ステップ140で作成したアンカーとリンク先識別子テーブルを検索し、京浜東北線の駅を指定するWebページへのURLを取得できる(ステップ136)。

【0133】以下同様にスタックのキーワードがなくなるまでステップ136からステップ140を繰り返すことで所望の情報を含むWebページにアクセスできる。前記ページ取得後は、TAGを解析して画面データを作成し表示手段に出力する(ステップ141)。

【0134】以上説明した方法により、所望のWebページへ順番にアクセスする必要がある場合でも。最初に

利用者がキーワードを音声で入力することで容易に所望の情報を取得することが可能である。

【0135】もちろんラティス解析や音声合成や関連情報選択ステップとの併用により、第1の実施例や第2の実施例で述べた効果を得ることが可能となる。

【0136】第4の実施例

本発明の第4の実施例を、図13と図14を用いて説明する。

【0137】第1、2、3の実施例と異なる点のみ説明するものとする。また、本実施例の効果を示せるように第1の実施例と同様の状況を想定する。

【0138】図14は表示方法の一例である。

【0139】図14の左側に示す表示領域20は、入力されたデータの種別を示す。この例では、上から順に、タスクの種類、路線名、駅名、施設名を示している。スロット21はデータ入力領域である。

【0140】なお、右側最上部はタスク名を示しており、利用者が入力すべき情報でないため、他のスロットと異なる方式で表示している。

【0141】1. ステップ151から156

ステップ151からステップ156は、入力された音声データと関連する情報をメモリに記憶する。

【0142】ここで、結果分類ステップ（ステップ153）は、認識結果に基づき対応表を調べ、入力されたデータを分類する。

【0143】結果表示ステップ（ステップ154）は、分類結果に基づきその種別に対応したスロットに表示データを提示する。

【0144】これにより利用者は、入力に対するシステムからのフィードバックを得ることが可能となり、認識エンジンが認識エラーを起こした場合でも、再度発話することにより、スロットのデータを更新することが可能となる。

【0145】以下のステップ157から161は、第2の実施例のステップ126から130と同様である。

【0146】以上により利用者は入力に対してフィードバックが得られ、間違っただけで認識された場合には、利用者は再度発話することにより訂正することができるので、スロットを指定する必要がなく利便性は向上する。

【0147】また、この方式では銀行のATMのようにシステム主導ではなく、利用者が任意のスロットを選択でき、利用者主導の対話が可能となる。

【0148】さらに、情報検索後に、メモリに記録されている前回の入力データをスロットに表示することにより、異なる条件で再度アクセスする場合に、利用者は変更すべき条件を容易に確認することができる。

【0149】第5の実施例

本発明の第5の実施例を説明する。

【0150】第1の実施例から第4の実施例まで、キーワードの入力手段として、音声認識手段のみ使用してい

るが、本実施例は、状況認識手段や、他の認識手段の結果をそのまま、あるいは、演算を施しキーワードを作成する。

【0151】本実施例は、図1に示す情報処理装置に、状況認識手段や1つ以上の認識用入力手段をバス9に付加して実現する。

【0152】状況認識手段としては、GPS、バーコードリーダー、無線TAG、赤外線通信、無線通信、有線通信などが利用できる。

【0153】認識用入力手段としては、カメラ、磁気センサ、加速度センサ、圧力センサ、ペンタブレット、節電位測定装置などが利用できる。認識用入力手段は、物理的に計測可能な信号に、予め決められた演算を行い、記憶手段3に記録されている辞書データと比較し、検索用のコードデータや制御用命令に変換する。

【0154】GPSを利用し緯度経度情報を入手する。

【0155】このデータをそのまま、あるいはデータベースをアクセスすることで地名や建造物、会社名に変換することで、場所や対象物に関するキーワードを得ることができる。これにより現在地や対象物に関する情報を入力しなくてもよい。

【0156】また、博物館がWebサーバと同様のサーバを有して、各展示物の情報を提供をする場合がある。

【0157】利用者が、目の前の展示物を指定する場合、バーコードリーダーを展示物のバーコードに近づけ、バーコードの値を対象物の識別コードとして用いることで、容易に対象物を特定することが可能となる。

【0158】あるいは赤外線送受信部を対象に向け、識別用データを得る。無線の有効範囲に立ち入ると、コネクションを設定する方法なども利用できる。

【0159】このような動作にあわせ、利用者は、音声認識手段で「製作年代」と発話することで、上記のコードと製作年代を示すコードを合わせて、サーバにURLを送信することで、利用者は容易に目の前の展示物の製作年代に関する情報をえることが可能となる。

【0160】さらに、利用者が、対象物の名前を知らない場合や、説明プレートの漢字が読めない場合などキーワードの発話ができないような状況でも、利用者は関連情報を取得可能となる。

【0161】画像認識、ジェスチャ認識、手書き文字認識、節電位認識などを併用することにより、どうしても音声認識が成功しない場合でも、所望のキーワードを入力することが可能となる。

【0162】以上説明したように、状況認識手段や、他の認識手段との併用により、音声認識で明示的にキーワードを入力しなくても、あるいは入力できないような状況でも、検索用のキーワードが準備でき、それゆえ利用者のキーワード入力に係る負荷をさらに低減したインタフェースの提供が可能となる。

【0163】

17

【発明の効果】本発明によれば、利用者は複数のキーワードの発話、あるいは文章を発話するだけで、ハイパーテキスト方式によってサーバから提供される複数の表示単位から構成される集合情報から、利用者が必要とする情報を検索する場合に、大幅に手間とアクセスまでの遅延時間を減少させることが可能となる。

【0164】利用者の望む情報のみ表示することで、スクロールなどの手間を省くことが可能となり、小型の情報処理装置での利便性が良くなる。

【0165】利用者主導での入力方式を提供することにより利便性を向上しつつ、入力時の手間を削減している。

【0166】音声以外の認識手段と組み合わせることで、入力の手間の削減や音声入力ができない場合でも所望の情報にアクセス可能となる。

【図面の簡単な説明】

【図1】本発明の第1の実施例を実現する情報処理装置のブロック図である。

【図2】サーバアクセス用の識別子の図である。

【図3】識別子の詳細説明図である。

【図4】第1の実施例に係るフローチャートである。

【図5】認識結果と属性の対応を示す図である。

【図6】識別子作成用のテンプレートを示す図である。

【図7】第1の実施例の変更例に係るフローチャートである。

18

*【図8】自然対話の例文とワードスポットによるキーワード認識を示す図である。

【図9】第2の実施例に係るフローチャートである。

【図10】サーバから取得した情報を示す図である。

【図11】応答文を示す図である。

【図12】第3の実施例に係るフローチャートである。

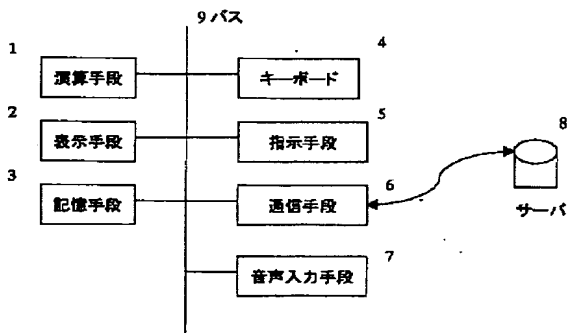
【図13】第4の実施例に係るフローチャートである。

【図14】キーワード入力用の画面を示す図である。

【符号の説明】

- 1 ……演算手段
- 2 ……表示手段
- 3 ……記憶手段
- 4 ……キーボード
- 5 ……指示手段
- 6 ……通信手段
- 7 ……音声入力手段
- 8 ……サーバ
- 9 ……バス
- 10 …URL
- 11 …サーバURL
- 12 …路線コード
- 13 …駅コード
- 14 …施設コード
- 20 …スロットの種別
- * 21 …スロットへの入力

【図1】



【図2】

<http://ekimae.toshiba.co.jp/spot/jibiki.htm?lineCode=001001&ekiCode=0000100&jibikiNo=3>

【図8】

「東京駅の京浜東北線近くのタクシー乗り場を教えてください」

【図11】

タクシー乗り場(4件)

JR山手線、丸ノ内南口周辺
JR山手線、八重洲中央口周辺
JR山手線、丸ノ内中央口周辺
丸ノ内線、出口2、3周辺

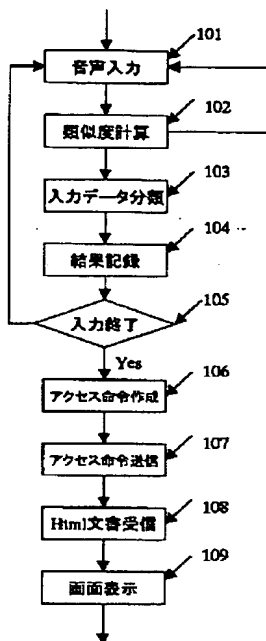
【図3】

【図5】

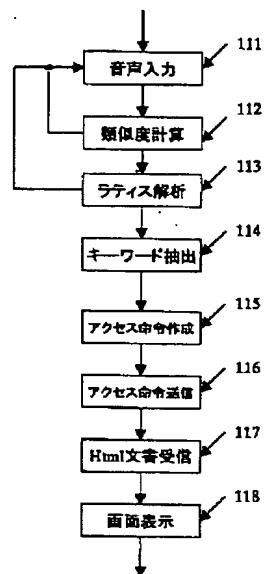
<http://ekimae.toshiba.co.jp/spot/jibiki.html> ← 11
lineCode=001001 ← 12
ekiCode=0000100 ← 13
jibikiNo=3 ← 14

11 けいひんとうほうく、京浜東北線、路線、001001
とうきょう、東京、駅、0000100
たぐしーのりば、タクシー乗り場、施設、3
けんさく、検索、命令

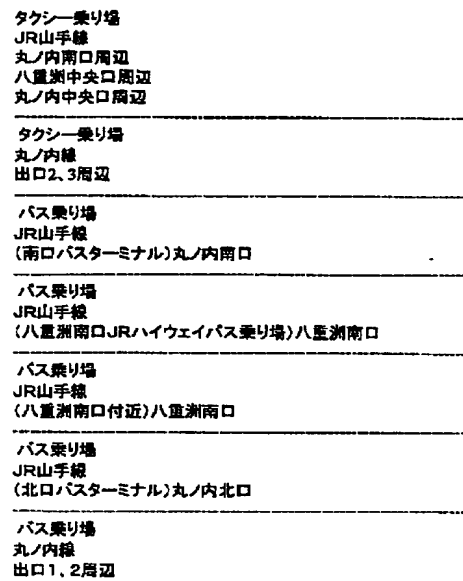
【図4】



【図7】



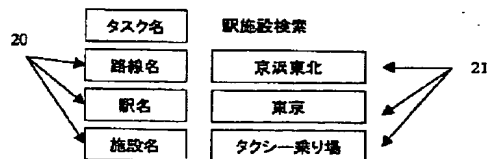
【図10】



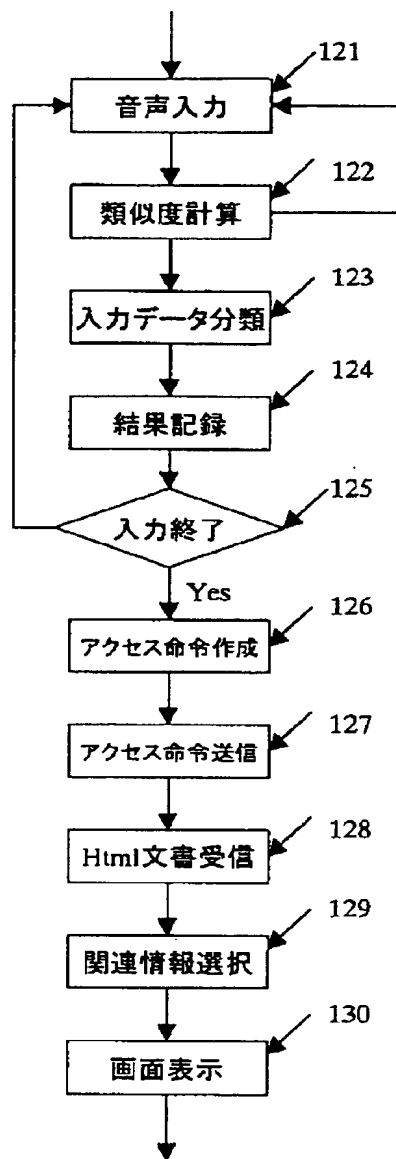
【図6】

<http://ckimae.toshiba.co.jp/spot/jibiki.html?lineCode=&ekiCode=&jibikiNo=>

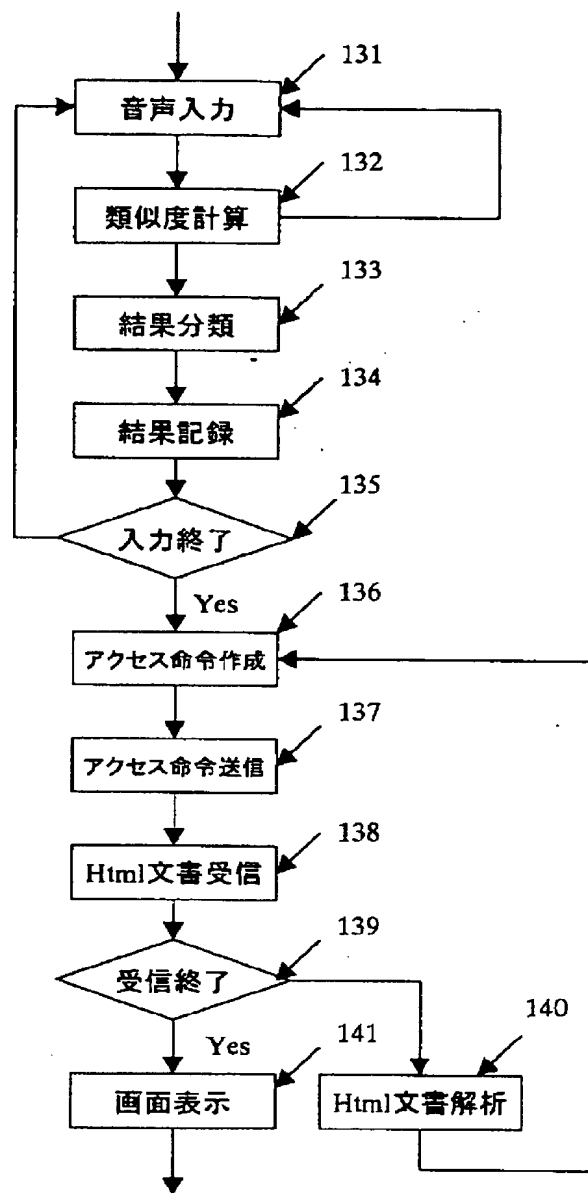
【図14】



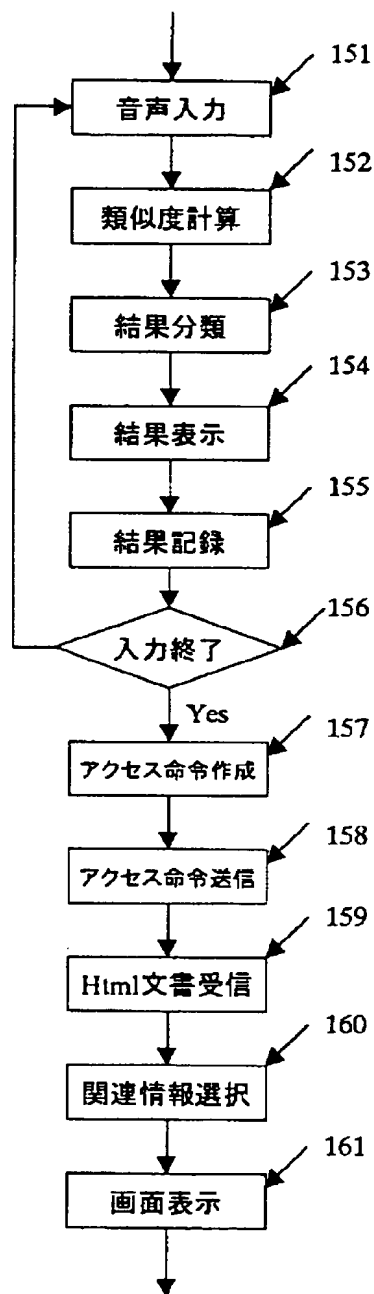
【図9】



【図12】



【図 13】



フロントページの続き

(51) Int. Cl. 7

G 1 0 L 15/18
15/24

識別記号

F I

G 1 0 L 3/00

テーマコード* (参考)

E

5 3 1 W

5 3 7 J

571Q

Fターム(参考) 5B075 KK07 ND20 PP07 PP12 PP30
PQ02 PQ42 QP10 UU40
5B082 GA08 GC04
5D015 AA04 BB01 HH03 HH06 HH13
HH16 LL01 LL05 LL06 LL09
LL12
5D045 AA04 AB01 AB04 AB26